

Przemysław SKUROWSKI  
Politechnika Śląska, Instytut Informatyki

## JPEG AND JPEG2000 COMPRESSION EVALUATION BASED ON S-CIELAB – HUMAN VISUAL SYSTEM MODEL

**Summary.** Both JPEG and JPEG2000 compression methods are based on human visual perception properties commonly there are used classical quality metrics. In the paper there is applied S-CIELAB filtering and  $\Delta E$  metric to evaluate quality of compression in JPEG standards measured as frequency weighted signal to noise ratio.

**Keywords:** image quality, human visual system, JPEG compression

## OCENA KOMPRESJI JPEG I JPEG2000 BAZUJĄCA NA MODELU LUDZKIEJ PERCEPCJI WIZUALNEJ – S-CIELAB

**Streszczenie.** Chociaż kompresja JPEG, jak i JPEG2000 bazują na pewnych cechach ludzkiej percepcji, do oceny jakości obrazów najczęściej wykorzystywane są proste metryki. W artykule wykorzystano do oceny jakości kompresji standardami JPEG ważony stosunek sygnału do szumu przy użyciu filtracji S-CIELAB i miary  $\Delta E$ .

**Słowa kluczowe:** jakość obrazów, model ludzkiego systemu widzenia, kompresja JPEG

### 1. Introduction

Image compression plays crucial role in modern networking due to large capacity of image data. The introduction of lossy JPEG compression in late 1980s (ISO10928 became official standard in 1992), as it significantly overcome the performance of lossless methods, allowed common using of relatively large images in network applications. The concept of JPEG compression is based on human limitations in visual performance. Nowadays, there is introduced its modern version - JPEG2000 [2, 6]. But despite the fact that both JPEG and

JPEG2000 are based on human factors, still quality metrics from conventional electrical engineering like SNR and MSE are mainly used. Recently there were developed human visual system models [4, 8] that approximate human perception of images which expanded classical observer model [1] to color spaces. The paper is an attempt to measure image reproducing quality in terms of human vision. Beginning with this paragraph for easy distinction between both compression standards basic JPEG will be named JPEG92.

## 2. Rough comparative review of JPEG92 and JPEG2000

Both JPEG92 and JPEG2000 in its lossy variant share common dataflow scheme although they differ in details. First the image is transformed from spatial to another domain (f.e. spectral). Then the image is quantized, linearized to form bitstream and coded using entropy coding. This scheme is depicted in figure 1. The main difference appears at the stage of bitstream forming where bitstreams are modeled and depending on the desired bitrate only the most significant band coefficients from wavelet domain are encoded

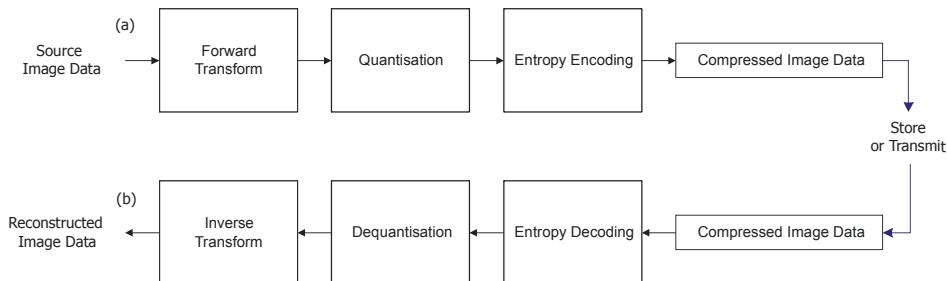


Fig. 1. Block diagrams of JPEG92/JPEG2000 encoder(a) and decoder (b)  
Rys. 1. Diagram blokowy kodera (a) i dekodera (b) JPEG92/JPEG2000

### 2.1. Transform coding of images

Lossy compression relies on removing information spatial redundancy (in human perception meaning). To achieve such an objective there are used transforms which allow to aggregate the most of image information (variance) within few coefficients that would be kept while all the rest will be erased as meaningless. The best concentration of variance within low coefficients is obtained by using Karhunen-Loeve Transform (KLT also known as Principal Component Decomposition) but this transform is inconvenient for use in practical applications. For practical purposes in JPEG92 there is Discrete Cosine Transform (DCT) used that well approximates KLT [5]. In case of JPEG2000 there is used discrete wavelet transform (DWT) instead of DCT. In fact both those transforms are used in spatio-frequency manner as block transforms applied to a fixed size tile of source image. Although for JPEG92 there are used tiles of 8x8 pixels whereas for JPEG2000 the tile size is a flexible parameter –

it is possible that the whole image is one tile. Significant improvement shared by both standards is subtracting value of half of range (probable mean) from each tile – in such a case only variance of image is processed.

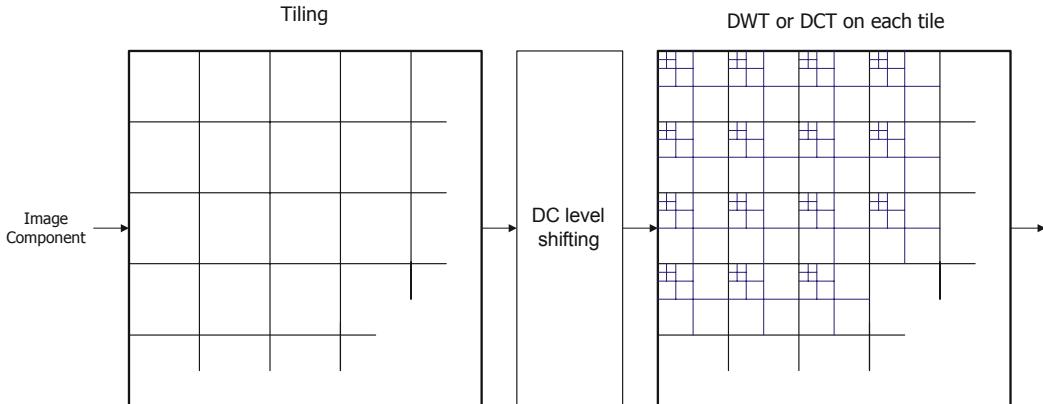


Fig. 2. Spatial partitioning of image into blocks

Rys. 2. Podział przestrzenny obrazu na bloki

## 2.2. Quantization

Loss of information happens at the stage of quantization. In case of JPEG92 there is *uniform scalar quantizer* (USQ) applied using quantization table according to formula:

$$F^Q(u,v) = \text{round}\left(\frac{F(u,v)}{Q(u,v)}\right) \quad (1)$$

where:  $F$  – spectral coefficients,  $Q$  – quantization table coefficients,  $F_Q$  – quantized values.

Basic version of JPEG2000 uses *deadzone scalar quantizer* that can be expressed as rounding towards zero (that is dead zone) using formula:

$$F^{Q_b}(u,v) = \text{sign}(F_b(u,v)) \text{floor}\left(\frac{|F_b(u,v)|}{\Delta_b}\right) \quad (2)$$

where:  $F_b$  – coefficients of transform at  $b$  band,  $F^{Q_b}$  – quantized coefficients of transform at  $b$  band,  $\Delta_b$  is quantization step for  $b$  band computed adaptively to dynamic range, bitdepth and energy of subband. Advanced version uses *Trellis Coded Quantizer* – TCQ (see [3]).

## 2.3. Entropy coding

The output stage of both JPEG coders applies two common lossless compression algorithms with some customizations to specific nature of compressed data. For JPEG92 it is *run-length coding* (*RLC*) combined with arithmetic or Huffman coding as option, meanwhile JPEG2000 uses MQ coder (approximated Elias coding as the first step and run length coding in the final step).

## 2.4. Bitrate control

Compression level is usually opposite demand to quality of image. It is measured by so called ‘bitrate’ expressed in average number of bits per pixel (bpp). The JPEG92 by default doesn’t control bitrate directly although it might be controlled by quality factor that is used to multiply quantization table coefficients – the higher quality the higher bit rate is obtained. For example the Independent Jpeg Group’s [9] codec `cjpeg` does quality control by quantization matrix scaling described in metacode below where `quality` is parameter supported by user.

```
Quality = 1..100
if (Quality < 50) { ScaleFactor = 5000 / Quality }
else { ScaleFactor = 200 - Quality * 2 }

for i in Matrix:
    NewQuantMatrix[i] = (StandardMatrix[i] * ScaleFactor + 50) / 100
```

The JPEG2000 controls bitrate directly by dropping the least significant subbands of wavelet transform of the image. This process is controlled by sophisticated algorithms [7].

## 3. Image evaluation based on human perception

### 3.1. Quality measures

Canonical image quality measures are simply derived from electrical engineering. Although they work well in machine vision, their efficiency is poor when they are used for approximation of human perception. This is caused by uniform treatment of all frequencies within image with no distinction between significant and almost invisible frequency components. There are two main families of such measure based on mean square error (MSE) and logarithmic signal to noise ratio (SNR). We can note them as formulas:

$$MSE(i_1, i_2) = \sum_{(m,n) \in i} (d(i_1(m,n), i_2(m,n)))^2, \quad (3)$$

$$SNR(i_1, i_2) = 10 \log \frac{\sum_{(m,n) \in i} i_1(m,n)^2}{\sum_{(m,n) \in i} (d(i_1(m,n), i_2(m,n)))^2} \quad (4)$$

where:  $d(\cdot)$  – a proper metric in color space ,  $m,n$  – image indices,  $i_1, i_2$  - compared images. To represent non uniform human perception it is necessary to take into account *contrast sensitivity function* as weighting function that describe human visual system as linear filter. [1] Image and difference of images in spectral domain are simply multiplied by CSF function.

Basing on convolution theorem it appears in spatial domain as convolution of image and spatial form of CSF. Weighted quality measures appear as:

$$WMSE(i_1, i_2) = \sum_{(m,n) \in i} (d(w * i_1(m,n), w * i_2(m,n)))^2, \quad (5)$$

$$WSNR(i_1, i_2) = 10 \log \frac{\sum_{(m,n) \in i} w * i_1(m,n)^2}{\sum_{(m,n) \in i} (d(w * i_1(m,n), w * i_2(m,n)))^2} \quad (6)$$

where:  $d(\cdot)$  – metric in color space,  $m, n$  – image indices,  $i_1, i_2$  – compared images,  $w$  – weighting function (contrast sensitivity function),  $*$  convolution operation.

### 3.2. S-CIELAB human visual model

Human vision is a complex process that is multi-stage, nonlinear and adaptive to viewing conditions. Classical color models like CIELAB and their color difference formulas were designed for color reception of relatively large areas so in such applications they work well. Simple predicting visible difference as sum of differences in CIELAB space for images having fine details usually poorly corresponds with human responses. Nowadays there are developed numerous *human visual system* (HVS) models that emulate treatment of image in the human visual pathway.

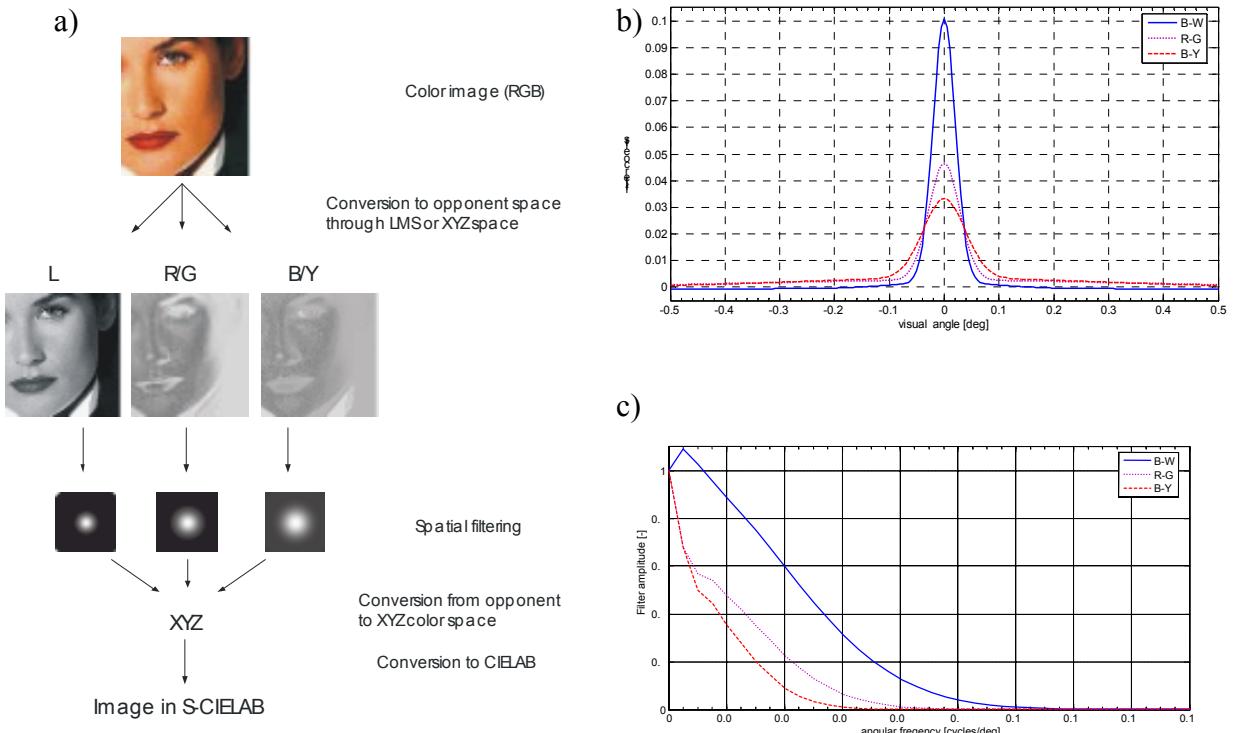


Fig. 3. S-CIELAB model: a) general dataflow; filters in b) spatial, c) spectral domain  
Rys. 3. Model S-CIELAB: a) ogólny schemat; filtry w postaci: b) przestrzennej, c) częstotliwościowej

One of such systems is S-CIELAB (spatial CIELab) [8] that is an enhancement of CIELAB model. It takes into account optical filtering and color conversion that take place in the visual pathway. Overall process is depicted in the fig. 3a, particular steps are performed as described below:

1. Conversion of image from RGB to opponent color (BW, RG, BY) space, This operation is done using intermediate LMS or XYZ (cone response) color space. Conversion itself is simple vector-matrix multiplication.
2. Spatial filtering performed as convolution with gaussian CSF filters with characteristics presented in fig. 3 b,c note it is expressed in cycles per degree.
3. Conversion from opponent color space to XYZ and then to CIELAB color space.

After such sequence of operations on two images we can compare them using any difference metries that applies in Lab color space. The most important one is Neugebauer distance expressed as Euclidian distance ( $\Delta E$ ) in perceptually uniform color space:

$$\Delta E = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2} \quad (7)$$

Using that distance we can compute or an error maps either cumulative error measures like SNR and MSE. Despite the convolution with CSF was performed in another color space we still call these measure as WSNR and WMSE respectively.

## 4. Experiment

### 4.1. Experimental setup

In the experiment there were used two well known codecs JPEG92 was compressed using `cjpeg` - Independent JPEG Group's codec [9] and JPEG2000 compressed using reference codec JJ2000 [10]. Both images were processed using SCIELAB filtering and there were computed error measures against original image using Neugebauer distance as error metric. In the experiments there were acquired (W)MSE and (W)SNR error measures versus bitrate. SCIELAB viewing conditions were set up to simulate observer watching image on monitor with standard resolution 72 DPI at 20 inches (approx. 0.5m) – so there are 25 pixels per degree in visual angle. The experiment was performed for six natural RGB color images commonly used for image testing with 8 bits/channel and joint entropies given in the table 1.

Table 1  
Test images and their joint entropies

airplane	baboon	couple	girl	lena	peppers
10.2932	12.7414	9.5126	10.8543	12.7647	12.5540

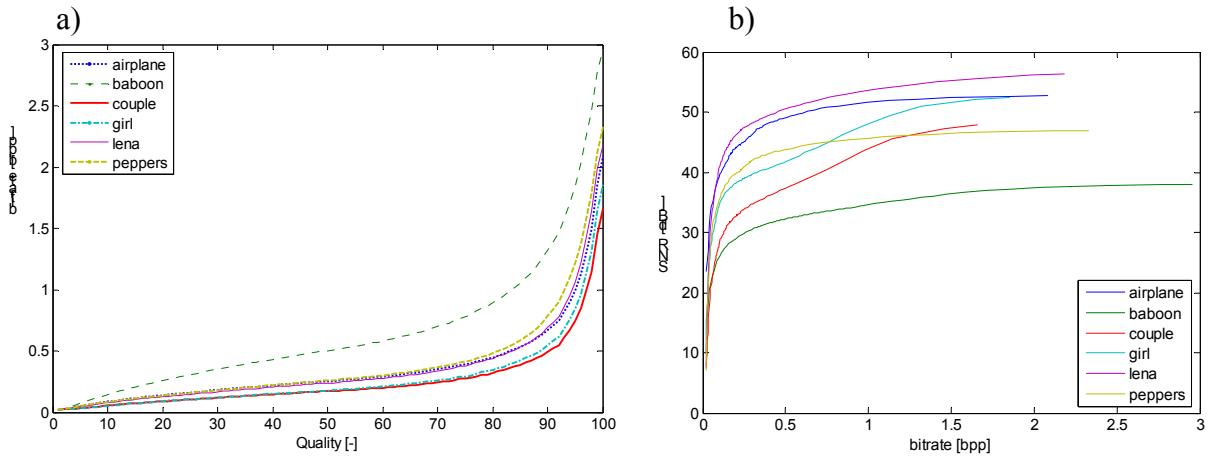


Fig. 4. JPEG92 a) Relation between Quality and bitrate , b) SNR versus bitrate  
Rys. 4. JPEG92: a) Bitrate w zależności od Quality, b) SNR w zależności od bitrate

It was impossible to control bitrate directly for JPEG92 moreover it was impossible to achieve the same range of bitrates for all images as their compressibility was different (see fig. 4b). Because `cjpeg` doesn't allow to control bitrate in the experiments there was used a trick – at first there was obtained output image for JPEG92 compression and then resulting bitrate (see fig 4a) was used as a parameter for JJ2000 codec. Testing procedure was performed for whole range of quality (1..100) for each image used in the experiment.

#### 4.2. Results

In the experiment there were obtained weighted and non-weighted SNR and MSE error measures. They are given at the figures below where on horizontal axis is Quality parameter – please note that we observe much more interesting regularity in shape of curve when it is presented as function of quality parameter than as a function of bitrate.

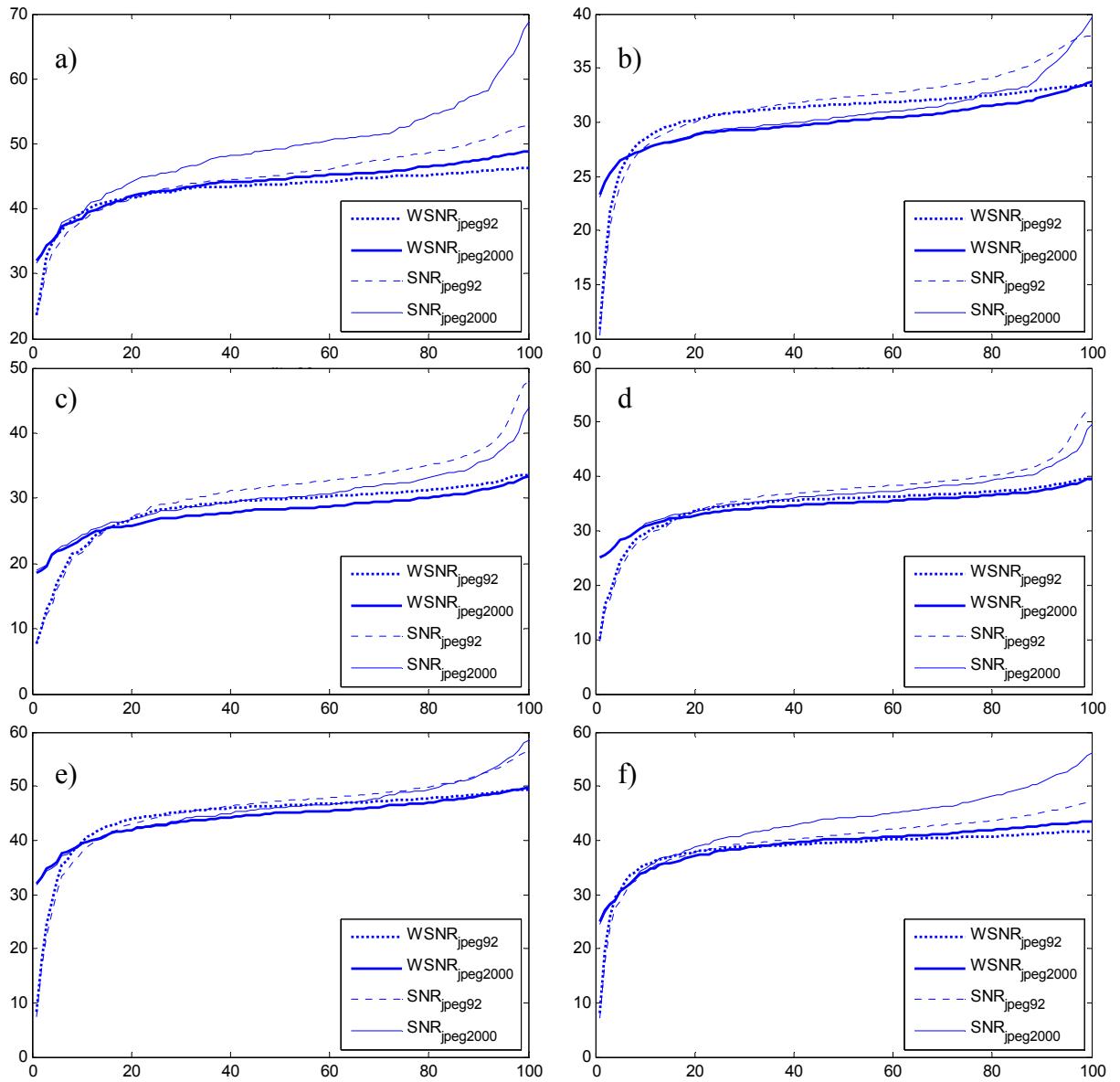


Fig. 5. (W)SNR measures in function of quality for JPEG92 and corresponding JPEG2000 for images: a) airplane, b) baboon, c) couple, d) girl, e) lena, f) peppers

Rys. 5. Miary jakości (W)SNR w funkcji parametru 'quality' dla JPEG92 i odpowiadającego JPEG2000 – obrazy: a) airplane, b) baboon, c) couple, d) girl, e) lena, f) peppers

#### 4.3. Observations

Generalized characteristics of observed results are shown at the fig. 6. It is easy to note a general observation that JPEG2000 has significantly better performance for low bitrates until point A at the fig. 6a. For moderate bitrates according to shown results JPEG92 has slightly better performance and for high bitrates (for B point) again JPEG2000 give a little bit better results. Accurate values of A and B points for specific images are placed in the table 2.

Table 2  
Intersection points for WSNR curves of JPEG92 and JPEG2000

Table 2

	airplane		baboon		couple		girl		lena		peppers	
	Quality	bpp	Quality	bpp	Quality	bpp	Quality	bpp	Quality	bpp	Quality	bpp
A	6	~0.06	7	~0.1	15	~0.07	17	~0.08	10	~0.08	5	~0.05
B	19	~0.13	98	~2.46	—	—	100	~1.86	98	~1.64	33	~0.19

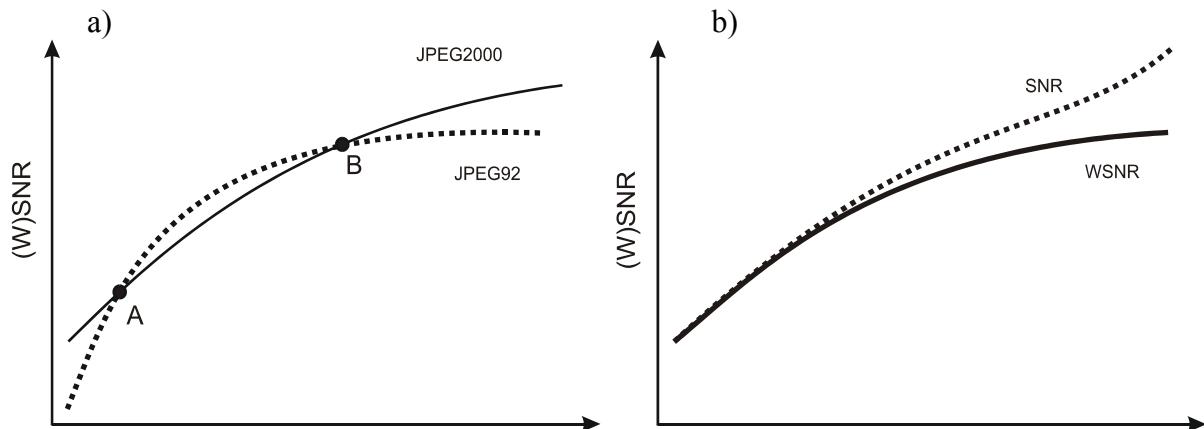


Fig. 6. Relation between Quality and bitrate for JPEG92  
Rys. 6. Bitrate w zależności od parametru Quality dla JPEG92

On the other hand there is interesting relation between classical SNR and WSNR. For images at low bitrate (poor quality) they conform each other but the higher bitrate/quality is measured the bigger difference is observed as it is roughly shown at fig. 6b. Such behavior of weighted measure corresponds well with human perception of quality of image – Quality set to 95 is commonly known to be maximal useful setting for jpeg.

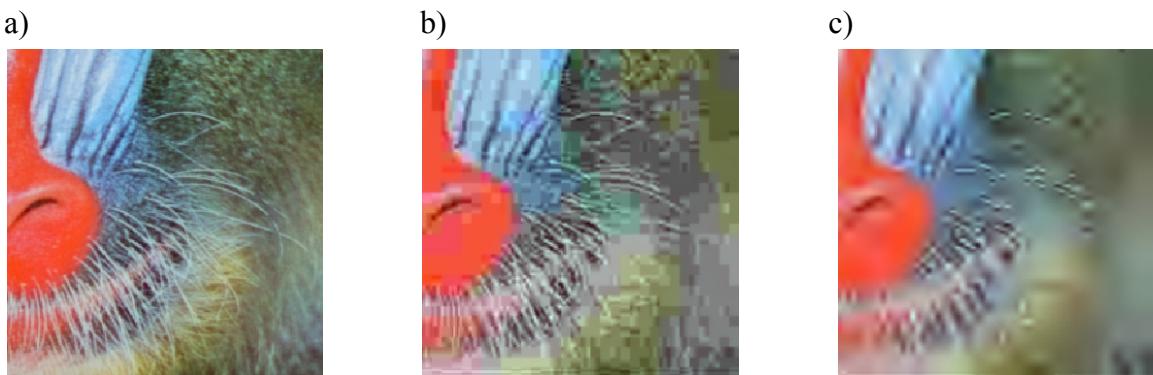


Fig. 7. Example of compression at low bitrate (0.07bpp): - a fragment of baboon image:  
a) original image, b) compressed using JPEG92, c) compressed using JPEG2000  
Rys. 7. Przykład kompresji przy niskim bitrate (0.07bpp) – fragment obrazu babon:  
a) obraz oryginalny, b) kompresowany JPEG92, c) kompresowany JPEG2000

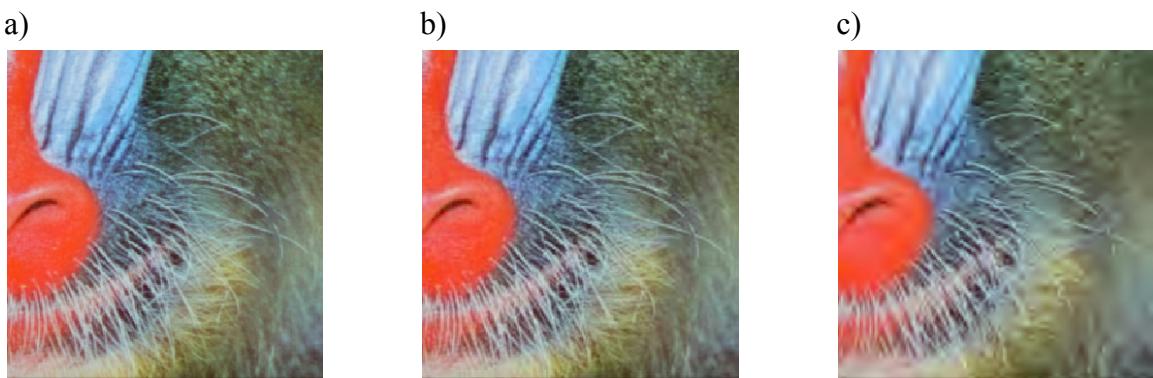


Fig. 8. Example of compression at medium bitrate (0.35bpp): - a fragment of baboon image:  
a) original image, b) compressed using JPEG92, c) compressed using JPEG2000

Rys. 8. Przykład kompresji przy średnim bitrate (0.35bpp) – fragment obrazu babon:  
a) obraz oryginalny, b) kompresowany JPEG92, c) kompresowany JPEG2000

## 5. Conclusions

Compression using JPEG2000 appears as a superior method for very low bitrates due to less blocking artifacts visibility (see fig. 7). Results observed in above experiment seems to confirm common observation that JPEG2000 sometimes has a little poorer visual efficiency then older JPEG92 – images compressed using JPEG2000 seem to be more ‘blurry’ (fig. 8). Explanation of that behavior might be in still poor quality of control algorithms which fit operating parameters of the JPEG2000 coder. Experimental results of applying S-CIELAB to measure quality of compression seem to be very promising method as it corresponds well with prior general observations of the compression effects. Although it is still necessary verify the results in formal way. In further work it is suggested to refer obtained results to some statistics of human responses like MOS (*mean opinion score*)

## BIBLIOGRAPHY

1. Campbell F. W.: The human eye as an optical filter. Proceedings of the IEEE, Vol 56 (6), pp. 1009÷1014, 1968.
2. ISO/IEC JTC1/SC29 WG1 (ITU-T SG8), FCD 15444-1, JPEG2000 Part I Final Committee Draft Version 1.0: JPEG2000 Image Coding System, 16 March 1997.
3. Marcellin M. W. et al.: An Overview of Quantization in JPEG2000. Signal Processing: Image Communications, Vol. 17 (1), pp. 73÷84, 2002.

4. Mantiuk R., Daly S., Myszkowski K., Seidel H.: Predicting Visible Differences in High Dynamic Range Images - Model and its Calibration. Proc. of Human Vision and Electronic Imaging X, IS&T/SPIE's 17th Annual Symposium on Electronic Imaging, pp. 204÷214, 2005.
5. Skarbek W. (red) et al.: Multimedia. Algorytmy i standardy kompresji. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1998.
6. Skodras A., Christopoulos C., Ebrahimi T.: The JPEG 2000 Still Image Compression Standard, IEEE Signal Processing Magazine, Vol. 18/5, pp. 36÷58, 2001.
7. Yu W., Sun F., Fritts J.E.: Efficient Rate Control for JPEG-2000. IEEE Trans. on Circuits and Systems for Video Technology, Vol. 16/5, pp. 577÷589, 2006.
8. Zhang X. M., Wandell B. A.: A spatial extension to CIELAB for digital color image reproduction. Proceedings of the SID Symposia , 1996.
9. <http://www.ijg.org>.
10. <http://jj2000.epfl.ch>.

Recenzent: Dr inż. Arkadiusz Sochan

Wpłynęło do Redakcji 3 kwietnia 2007 r.

## Omówienie

Artykuł jest próbą zastosowania miar błędu uwzględniających ludzki czynnik w ocenie do oceny jakości stratnej kompresji obrazów. Korzystając z modelu S-CIELAB porównano klasyczną metodę JPEG z jej nowszą odmianą JPEG2000. Zestaw sześciu testowych obrazów został skompresowany odpowiednimi koderami przy zmiennym parametrze reprezentującym jakość kompresowanego obrazu, następnie obrazy zostały zdekompresowane i porównane z obrazami źródłowymi. W ocenie jakości obrazu wykorzystano jako miarę błędu ważony logarytmiczny stosunek sygnału szumu (WSNR) oraz prostą miarę SNR, gdzie miarą bliskości kolorów dla poszczególnych pikseli była odległość Neugebauera, czyli metryka euklidesowa w przestrzeni jednorodnej percepcyjnie. W artykule porównano wyniki uzyskane miarą WSNR z klasycznym błędem SNR. Artykuł kończy wskazanie kierunku dalszych badań.

**Adres**

Przemysław Skurowski: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16,  
44-100 Gliwice, Polska, Przemyslaw.Skurowski@polsl.pl .